



①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ **Offenlegungsschrift**
⑩ **DE 196 24 977 A 1**

⑤① Int. Cl.⁸:
G 06 K 9/48
B 07 C 3/12
G 06 K 9/32

②① Aktenzeichen: 196 24 977.5
②② Anmeldetag: 22. 6. 96
④③ Offenlegungstag: 2. 1. 98

DE 196 24 977 A 1

⑦① Anmelder:
Siemens AG, 80333 München, DE

⑦② Erfinder:
Hänsel, Karl-Günther, 78464 Konstanz, DE;
Rosenbaum, Walter, Dr., Paris, FR

⑤⑥ Entgegenhaltungen:
DE 1 95 31 392 C1
DE 43 04 082 A1
US 49 92 649 A
IBM-Prospekt »M.A.R.S. - IBM Automatic Mail
Address Recognition System«, May 1995;
BUHCARD/RUSSAU: »Die Post geht ab«, TR
Transfer 35/94, S. 98/99;

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ Verfahren zur Verarbeitung von Postsachen

⑤⑦ Bei einem Verfahren zur Bearbeitung von Sendungen in einem automatischen Adreßlesesystem, wobei ein Abbild der Adreßinformation aufweisenden Oberfläche jeder Sendung gewonnen und einer ersten automatischen Auswertung zugeführt wird, und nicht vollständig ausgewertete Adreßinformation einer ersten Video-Codierung zur Auswertung zugeführt wird, ist vorgesehen, daß die Adreßinformation derjenigen Abbilder, die bei der Video-Codierung nicht vollständig ausgewertet worden sind, unter Verwendung der Ergebnisse der Video-Codierung einer weiteren automatischen Auswertung zugeführt werden.

DE 196 24 977 A 1

Beschreibung

Die Erfindung betrifft ein Verfahren gemäß dem Oberbegriff der unabhängigen Patentansprüche. Im weitesten Sinne bezieht sich die Erfindung auf das Gebiet automatischer Briefverarbeitung und insbesondere auf Systeme, bei denen ein automatisches Adreßleseverfahren durch Videocodierung ergänzt und verbessert wird.

Systeme zum automatischen Lesen von Adressen (OCR) sind auf dem Gebiet der Briefbearbeitung wohlbekannt und z. B. in der DE 195 31 392 beschrieben. Mit modernen OCR-Briefsortieranlagen sind Bearbeitungsraten von 10 Briefen pro Sekunde, d. h. 36.000 Briefen pro Stunde und mehr erreichbar. Die Erkennungszuverlässigkeit variiert jedoch stark mit der Schriftart und Gesamtqualität der auf die Oberfläche der Briefe aufgetragenen Adreßinformation.

Im Falle einer erfolgreichen Erkennung kann der betreffende Brief mit einem maschinenlesbaren Barcode versehen werden. Dieser Barcode erlaubt eine weitere mechanische Bearbeitung bis zu einer beliebigen gewünschten Sortierordnung.

Insbesondere ermöglicht die Verwendung von Barcodes eine Sortierung der Briefe bis zur Sortierebene des Postgangs, bei der Briefe entsprechend der Reihenfolge ihrer Verteilung durch den Zusteller sortiert werden.

Da die Erkennungsraten der automatischen Lesesysteme sehr variieren, ist es notwendig, diese durch verschiedene Formen von manueller Intervention zu unterstützen. Die einfachste Intervention ist Zurückweisung nicht automatisch lesbarer Briefe, und die Durchführung eines Handsortierverfahrens. Allerdings sind die dadurch entstehenden Kosten mit zunehmenden Arbeitskosten unökonomisch hoch. Hinzu kommt, daß solche handsortierte Post auch zu späteren Zeitpunkten nicht mehr ohne weiteres mechanisch weitersortiert werden kann, so daß zwei voneinander getrennte Ströme von Sendungen erzeugt werden, die wiederum manuell zu einem bestimmten Zeitpunkt verschmolzen werden müssen.

Um diese Nachteile manueller Sortierung von OCR-zurückgewiesenen Sendungen zu vermeiden, sind verschiedene Verfahren für eine manuelle Codierung von Poststücken entwickelt worden. Alle diese Verfahren verwenden Interventionen durch Operatoren, um Barcodes auf die Sendungen auf eine Weise aufzubringen, die konsistent ist mit dem Erfordernis, eine mechanische Sortierung durch die gleichen Maschinen vorzunehmen, die OCR-gelesene und barcodierte Post bearbeiten.

Ein weiteres Verfahren um zurückgewiesene Poststücke zu codieren verwendet sogenannte manuelle Codierplätze. An diesen manuellen Codierplätzen werden die Sendungen physikalisch hintereinander einem Operator vorgeführt, wobei der Operator für jedes dieser Sendungen so viel Information codiert wie notwendig ist, um den Bestimmungsort eindeutig zu identifizieren. Dabei wird die eingegebene Adresse mittels eines Zeichnisses in einen Sortierbarcode umgewandelt, der auf die Sendung aufgebracht wird. Die codierten Sendungen werden dann mittels Barcodesortierer (BCS), die identisch mit OCR-geeigneten BCS sind, weiterbearbeitet. Solche manuellen Codierplätze wurden als erste von der US-Post und der Royal Mail in den 70er Jahren eingeführt. Die Hauptnachteile solcher Vorrichtungen bestehen in der notwendigen Entfernung von Sendungen aus dem OCR-Sendungsstrom und ergonomischen Schwierigkeiten für den Operator beim Erkennen der

ihm vorbeigeführten Sendungen.

Der nächste Fortschritt in der Behandlung von OCR-zurückgewiesener Sendungen war die Entwicklung von On-line-Videocodiersystemen (OVS). In einem OVS wird ein Videobild der Sendung einem Operator zum Codieren vorgeführt, statt der physikalischen Sendung bei den manuellen Codierplätzen. Das Videobild wird dem Operator gezeigt während die physikalische Sendung in Verzögerungsstrecken gehalten wird. In diesen Verzögerungsstrecken wird normalerweise die Sendung in Bewegung gehalten für eine Zeitdauer, die ausreicht, damit der OVS-Operator die notwendige Sortierinformation für das betreffende Bild eingibt. Die üblichen Verzögerungsstrecken erlauben eine Verzögerung zwischen 10 und 30 Sekunden. Je länger die Verzögerungsstrecke, desto größer sind die Kosten sowie die Erfordernisse für Wartung und die physische Größe der Anlage.

Das Hauptproblem bei der Verwendung von OVS ist, daß die zur Verfügung stehende Zeit lediglich ausreicht für eine sorgfältige Eingabe der Postleitzahl (ZIP) oder des Postcodes (PC), es sei denn, daß unpraktikabel lange Verzögerungsstrecken verwendet werden.

Daher wurden spezielle Codierv Verfahren entwickelt, um die notwendige On-line-Verzögerungszeit möglichst gering zu halten.

Um die Codierproduktivität zu erhöhen und/oder die Angabe von sämtlichen Adreßelementen, d. h. ZIP/PC, Straße/Postfach, Adressat/Postfach, Adressat/Firma, zu ermöglichen, sind im Stand der Technik verschiedene Methoden entwickelt worden. Im wesentlichen sind dies:

Vorschau-Codierung

Bei der Vorschau-Codierung erfolgt ein simultaner Display der Bilder von zwei Sendungen erfolgt; eines über dem anderen. Hierbei ist das niedrigere Bild das aktive, d. h. dessen Daten codiert werden. Nach einem geeigneten Training ist es Operatoren möglich die Information auf dem unteren Bild zu codieren, während sie bereits die Adreßinformation vom oberen Bild aufnehmen. Das obere Bild wird anschließend aktiv und der Prozeß wird fortgesetzt. Mit der Vorschau-Codierung ist es möglich, durch eine vollständige Überlappung der kognitiven und motorischen Funktionen beim Codieren aufeinanderfolgender Abbilder die Operator-Produktivität zu verdoppeln.

Extraktionscodierung

Da bei den praktisch erreichbaren On-line-Verzögerungszeiten lediglich die ZIP/PC Adreßelemente verläßlich vom Operator eingegeben werden können, werden beim Extraktionscodieren bestimmte Schlüsselbestandteile der sich auf die Straße beziehenden Adreßbestandteile eingegeben. Üblicherweise basiert die Extraktionscodierung auf speziell entwickelten Regeln, bei denen ein Code fester Länge als Zugangsschlüssel zu einem Adreßverzeichnis verwendet wird. Zum Beispiel verwendet die Royal Mail eine Extraktionsformel, die auf den ersten drei und den letzten zwei Buchstaben basiert. Dabei müssen Sonderregeln vom Operator auswendig gelernt werden, um überflüssige Adreßinformationen zu vermeiden und bestimmte unterscheidende Merkmale wie z. B. Richtungen zu berücksichtigen, z. B. Ost, West oder Kategorien z. B. Street, Lane, Road.

Extraktionscodierung besitzt trotz ihrer gewissen Ef-

ektivität einige größere Nachteile; insbesondere komplexe Extraktionsregeln, die häufig die Berücksichtigung des Endes eines Straßennamens erfordern, während diese Bestandteile üblicherweise am unklarsten geschrieben sind. Außerdem findet man eine signifikant hohe Rate von nicht eindeutigen Extraktionen bei denen mehrere Einträge in einem Verzeichnis dem Extraktionscode entsprechen, so daß keine eindeutige Sortierentscheidung getroffen werden kann. Desweiteren ist zu berücksichtigen, daß die Eingabeproduktivität der Operatoren vermindert wird, sobald vom Operator Entscheidungen getroffen werden müssen, statt einer einfachen repetitiven Tastatureingabe.

Komplettierungscodierung

Bei der Komplettierungscodierung wird im Kontrast zur Extraktionscodierung für jede zu kodierende Adresse eine variable Eingabe vorgenommen. Im wesentlichen wird bei der Adreßeingabe ein Abgleich mit dem Adreßverzeichnis vorgenommen, bis eine eindeutige Übereinstimmung erreicht ist. Durch die Darstellung des Rests der Adresse sobald eine eindeutige Teilübereinstimmung identifiziert worden ist, wird ein Beschleunigungseffekt erreicht. Probleme treten bei dieser Technik jedoch dadurch auf, daß dem Operator ein Eingabestop-Signal zugeführt werden muß und eine Darstellung des identifizierten Adreßrests notwendig ist, was zu einer verminderten Eingabeproduktivität führt und eine Vorausschaukodierung unmöglich macht.

Operatorassistierte OCR-Technik

Zur Erhöhung der On-line zur verarbeitenden Adreßinformation hat die US Post mit Operator assistierten OCR-Techniken experimentiert. Hierbei wird zur Erhöhung der Effektivität der Teil des Adreßbildes hervorgehoben, bei dem die OCR-Erkennung versagt hat. Da die Operatoren bei der Entzifferung von fehlenden Buchstaben langsam sind und zum Teil auch komplexe Erkennungsfehler, z. B. Segmentierungsprobleme auftreten, ist bei diesem Verfahren die Operatorproduktivität häufig niedriger als bei einfachen Wiedereingabe der betreffenden Adresse.

Off-line Codierung

Da bei keiner der oben erwähnten Codierungstechniken eine ausreichend hohe Produktivität bei reiner On-line-Codierung erreicht wird, wurde kürzlich ein off-line Codiersystem eingeführt, wie in der US PS 49 92 649 beschrieben. Bei diesem System werden Sendungen mit nicht erkannten Adressen mit einer zusätzlichen Information, einer tracking identification (TID) versehen. Die nicht erkannten Sendungen werden extern gespeichert, während die Bilder dieser Sendungen Operatoren zum Codieren präsentiert werden wobei keine zeitlichen Einschränkungen vorliegen. Anschließend werden die Sendungen TID-Lesegeräten vorgeführt. Die TID wird mit der eingegebenen Adreßinformation verknüpft. Darauf basierend kann auch eine übliche Barcode-Sortierinformation auf die Sendung aufgebracht werden, so daß die betreffende Sendung wie üblicherweise OCR-gelesene Sendungen bearbeitet werden können. Obwohl das off-line-Video-Codierverfahren eine effektive Methode zur Codierung sämtlicher Adreßbestandteile darstellt, werden zusätzlich Kapazitäten für die Weiterverarbeitung mit nichtgelesenen Adressen vorgesehen

Sendungen und eine entsprechend komplexe Logistik verlangt.

Aufgabe der vorliegenden Erfindung ist es, eine möglichst hohe on-line-Codierung von Briefsendungen und eine möglichst geringe Rate von Sendungen mit nicht erfolgreich gelesenen Adressen zu erreichen. Erfindungsgemäß wird diese Aufgabe durch die Merkmale der unabhängigen Patentansprüche gelöst.

Die Erfindung ermöglicht eine verbesserte Integration von automatischen Lesesystemen und Video-Codierung. Weiter ermöglicht die Erfindung die effektive Verwendung einer Extraktionscodierung in integrierten automatischen und Video-Codiersystemen, insbesondere auch eine Vereinfachung des Entscheidungsproblems für die Operatoren bei der Adreßcodierung. Ein weiterer Vorteil des erfindungsgemäßen Verfahrens ist, daß zusätzliche Sortierinformationen, z. B. solche, die sich auf die Namenszeile der Adresse beziehen, effektiv ausgewertet werden können. Ein weiterer Vorteil der Erfindung besteht darin, daß Inkonsistenzen, die sich z. B. bei der Extraktionscodierung oder durch Fehler bei der Beschriftung von seitens des Absenders ergeben, effektiv behandelt werden können. Das Verfahren ermöglicht weiter eine einfache Integration von on-line und off-line-Codierung sowie des Vorschau-Codierungsverfahrens.

Vorteilhafte Ausführungsformen der Erfindung sind den abhängigen Ansprüchen und der Beschreibung zu entnehmen.

Im folgenden wird die Erfindung anhand von Zeichnungen beispielhaft dargestellt. Im einzelnen zeigen:

Fig. 1 eine schematische Darstellung einer Vorrichtung zur Ausführung des Verfahrens,

Fig. 2 und 3 eine Übersicht über den Datenfluß gemäß der Erfindung.

Fig. 1 zeigt eine schematische Darstellung einer Briefverteilanlage mit der das erfindungsgemäße Verfahren ausgeführt werden kann. Der OCR-Briefsortierer 100 besteht aus einer Zuführeinrichtung 110, die sukzessive Sendungen aus einem Magazin 111 abzieht und mit ca. 10 Sendungen pro Sekunde zu einem hochauflösenden Videoscanner 120 transportiert. Anschließend werden die Sendungen in einer Verzögerungsstrecke 121 transportiert. Die Sendungen weisen üblicherweise auf ihrer Oberfläche Adreßinformationen auf. Im OCR-Prozessor 130 erfolgt eine Auswertung der Adreßinformationen der Abbilder der Sendungen die vom Videoscanner 120 gewonnen wurden. Im Fall einer vollständigen Auswertung wird ein Barcodedrucker 150 angesteuert und die Sendung wird mit einem entsprechenden Barcode für die anschließende Sortierung in Sortierfächer 160 versehen. Der OCR-Prozessor 130 besteht aus einem oder mehreren Mikroprozessoren 131 mit assoziiertem Speicher 132 um Abbilder der Sendungen zu speichern. Weiterhin beinhaltet der OCR-Prozessor ein Adreßverzeichnis 134 mit ZIP-Codes, Städtenamen und Straßennamen und evtl. weiterer adreßbezogener Information. Der bei der Auswertung der Adreßinformation aufweisenden Abbilder erfolgt eine merkmalsgesteuerte Reduktion der aus dem Adreßverzeichnis gewonnenen Eintragung, so daß eine Art Wörterbuch erzeugt wird. In einzelnen Eintragungen werden dabei Glaubwürdigkeiten zugeordnet, so daß bei der Auswertung eine Anzahl von Daten von richtig erkannten Adressen erzeugt werden. Die Vorrichtung enthält weiterhin einen Imagecontroller 170 sowie eine Anzahl Video-Codierplätze 200, die mit dem Imagecontroller 170 direkt oder durch ein lokales Netzwerk

(LAN) 171 verbunden sind. Falls die OCR-Auswertung eines Abbildes nicht, d. h. vollständig erfolgreich war, wird dieses Bild vom OCR-Prozessor 130 zum Image-Controller 170 transferiert, der einerseits den TID Barcode-Drucker 151 steuert und andererseits das entsprechende Bild zu einer der Video-Codierplätze 200 sendet. Der TID Barcode-Printer 151 bringt auf die entsprechende Sendung einen Identifikationscode TID auf, der es zu einem späteren Zeitpunkt ermöglicht, die ausgewertete Adreßinformation mit der physikalischen Sendung zu knüpfen. Die Auswertung der Abbilder erfolgt in diesem Fall vorzugsweise off-line, obwohl grundsätzlich bei einer genügend langen Verzögerungszeit auch eine on-line-Auswertung durch Video-Codierung möglich ist. Im letzteren Fall kann der TID auch zu einem späteren Zeitpunkt, d. h. wenn die Video-Codierung innerhalb einer bestimmten vorgegebenen Zeit nicht zu einer vollständigen Auswertung geführt hat, auf die Sendungen aufgebracht werden.

Zur Ausführung des weiteren genauer dargestellten Verfahrens ist der Image-Controller 170 derart ausgebildet, daß von der Video-Codierung nicht vollständig ausgewertete Adreßinformationen unter Verwendung der Ergebnisse der Video-Codierung im OCR-Prozessor einer weiteren automatischen Adreßauswertung zugeführt werden.

Fig. 2 und 3 zeigen Datenflußdiagramme. Vorzugsweise arbeiten die Operatoren mit geteilten Displays 210 von denen der obere eine Vorschau ermöglicht und der untere üblicherweise aktiv ist. Vom Operator eingegebene Daten werden in der Promptzeile 211 dargestellt. Zum Videocodieren gelangen solche Abbilder von Sendungen, die nicht vollständig automatisch ausgewertet wurden 220. Beim Beispiel der Fig. 2 erfolgt eine Eingabe eines ZIP-Codes "4431", eines Extraktionscodes für die Straße "Hell" sowie die Hausnummer "8". Statt eines zweigeteilten Displays sind auch andere Formen möglich. Diese Eingabeinformationen werden benutzt, um im Adressenverzeichnis 134 passende Einträge zu identifizieren. Für den Fall, daß eine eindeutige Zuordnung zwischen der eingegebenen Information und einem Eintrag im Adressenverzeichnis 134 gefunden wurde, ist damit grundsätzlich eine vollständige Auswertung der Adreßinformation des betreffenden Abbildes erfolgt. Allerdings ist für einen gewissen Prozentsatz von Eingaben eine solche Eindeutigkeit nicht erreichbar, da eine Anzahl verschiedener Adreßeinträge der codierten Adreßinformation entsprechen.

Gemäß Fig. 3 erfolgt eine Entscheidung 300, ob eine vollständige Auswertung der Adreßinformation eines Abbildes bei der Video-Codierung erfolgt ist. Ist die Entscheidung positiv (Yes) kann die entsprechende Sendung entweder mit einem Barcode versehen werden, wenn die Verzögerungszeit ausgereicht hat, um eine On-line-Video-Codierung durchzuführen oder es wird eine entsprechende Verknüpfung der TID und eine darauf basierende Barcodierung vorgenommen. In jedem Fall kann die entsprechende Sendung mit üblichen Mitteln weiter sortiert werden. Bei einer negativen Entscheidung (No) erfolgt erfindungsgemäß eine weitere automatische Auswertung unter Verwendung der Ergebnisse der Video-Codierung, d. h. bei dieser weiteren automatischen Auswertung stehen dem OCR-Prozessor neben den auf der Abbild dargestellten Informationen die durch die Video-Codierung gewonnenen Informationen zur Verfügung; im obigen Beispiel das Tripel "4432", "Hell", "8". Dies ist in Fig. 3 symbolisch durch den Inhalt des Kreises 310 zum Ausdruck gebracht. An-

schließend erfolgt eine Entscheidung 320, ob nunmehr eine vollständige Auswertung des betreffenden Bildes stattgefunden hat. Im positiven Fall (Yes) wird die entsprechende Information zum weiteren Sortieren der Sendung verwendet analog wie nach einer positiven Entscheidung beim Entscheidungspunkt 300. Ist die Entscheidung negativ (No) erfolgt eine weitere Video-Codierung unter Verwendung der Ergebnisse der weiteren automatischen Auswertung. Vorzugsweise werden dabei dem Operator eine Anzahl von Alternativen zur Selektion vorgeführt, aus denen eine Selektion vorzunehmen ist.

Vorzugsweise werden die beiden letzten Schritte des Verfahrens nämlich die weitere automatische Auswertung sowie die weitere Video-Codierung nicht mehr on-line sondern off-line erfolgen, weil die zur Verfügung stehende Verzögerungszeit zu gering ist für eine on-line-Ausführung.

Das erfindungsgemäße Verfahren beinhaltet demnach drei zusammenwirkende Phasen. Diese sind:

1. Dateneingabephase durch Video-Codierung, bei der vorzugsweise mit einem einfachen Extraktionscode eine Codierung bestimmter Teile der Adreßinformation erfolgt. Üblicherweise wird hierbei die ZIP/PC Information sowie ein größerer Teil der weiteren Adreßinformation vollständig ausgewertet oder es erfolgt bei fehlenden ZIP/PC eine Extraktion des Ortsnamens. Dieser Phase vorgeschaltet war bereits eine erste automatische Auswertung der Adreßinformation.

Vorzugsweise erfolgt eine Eingabe mit einem geteilten Display. Für Vereinfachung der Eingabe wird ein einfacher Extraktionscode verwendet, z. B. ein 4stelliger Postcode, die ersten vier Alpha-Zeichen des Straßennamens und die Ziffern der Hausnummern der betreffenden Adresse. Bei dieser Extraktionscodierung ist eine Anpassung an die jeweiligen Postkonventionen ohne weiteres möglich. So kann z. B. die Anzahl der ersten Buchstaben variiert werden. Vorzugsweise wird der Postcode vom Operator nur dann eingegeben, wenn die OCR-Auswertung überhaupt kein Ergebnis gebracht hat. Die Mehrzahl der Sendungen wird daher die Eingabe von Straßenangaben betreffen. Vorzugsweise kann auch eine Strukturierung der Video-Codierung derart erfolgen, daß eine Gruppe von Operatoren Postcode und Straßenangaben eingibt, während eine andere Gruppe lediglich Straßenangaben eingibt. Da ein bestimmter Prozentsatz der Post heutzutage Postfachangaben enthält, wird vorzugsweise eine geeignete Taste mit einem Postfachschlüssel belegt die gegebenenfalls vom Operator gedrückt wird, wobei anschließend die Eingabe der Postfachnummer folgt. Für Firmenadressen, bei denen Straßen- oder Postfachangaben fehlen, ist auch die Eingabe der Firmenzeile möglich.

2. Weitere automatische Auswertung unter Verwendung der Informationen, die bei der vorhergehenden Phase durch Video-Codierung eingegeben worden sind. Die zusätzliche Information erhöht die Wahrscheinlichkeit einer vollständigen Auswertung bei der weiteren automatischen Auswertung.

Bei nicht vollständig in der vorhergehenden Phase ausgewertete Adreßinformationen folgt eine weitere automatische Auswertung. Wird mit einem Ex-

traktionscode gearbeitet, so liegen für diese zwei oder mehr Eintragungen im Adreßverzeichnis vor. Bei geeigneter Extraktionscodierung sind dabei in über 90% der Fälle lediglich zwei Adreßeinträge vorhanden.

3. Weitere Video-Codierung derjenigen Abbilder, für die bei den vorhergehenden Phasen keine vollständige Auswertung möglich war.

Vorzugsweise werden Abbilder der nicht ausgewerteten Adressen dem Operator präsentiert, zusammen mit den Ergebnissen der vorhergehenden Video-Codierung und der automatischen Auswertung der zweiten Phase zwei. Vorzugsweise hat der Operator hierbei lediglich eine Selektion aus vorgegebenen Alternativen vorzunehmen.

Bei der weiteren automatischen Auswertung liegt nunmehr zusätzliche Kontextinformationen vor und zwar ist die Anzahl der zu berücksichtigten Adreßeinträge beschränkt durch die Anzahl der aus der Extraktionscodierung erhaltenen Kandidaten. Es kann davon ausgegangen werden, daß unter diesen Kandidaten die richtige Adresse zu finden ist. Zusätzlich ist üblicherweise die Hausnummer bekannt.

Bei der weiteren Video-Codierung solcher Adreßinformationen, die in den vorhergehenden Phasen nicht vollständig ausgewertet wurden, werden vorzugsweise nicht eindeutige Ergebnisse der Extraktionscodierung oder zusätzliche sortierrelevante Informationen, die sich auf der Adressatenlinie befinden, bearbeitet. Vorzugsweise werden dem Operator sukzessive Bilder der Sendungsoberfläche präsentiert, wobei in einem Fenster die Auswertungsalternativen der Adresse oder des Adressaten dargestellt sind. Die Auswahl kann entweder durch Tastatureingabe einer Auswahlnummer, per Maus oder über einen Sprachprozessor erfolgen.

Bei einer bevorzugten Ausführungsform der Erfindung erfolgt eine Auswertung eines ersten Bestandteils der Adreßinformation und eine Auswertung eines zweiten Bestandteils der Adreßinformation und eine Prüfung der Ergebnisse dieser Auswertungen in Hinblick auf gegenseitige Konsistenz. Der erste Bestandteil der Adreßinformation kann insbesondere der ZIP- oder Postcode sein, der zweite eine Straßenangabe oder eine Hausnummer. Inkonsistenzen zwischen diesen beiden Bestandteilen können z. B. durch Lesefehler oder durch eine Falschangabe des ZIP- oder Postcodes vorkommen. Bei einer Falschangabe werden ausgehend von den ersten drei oder vier Ziffern des ZIP oder Postcodes eine Anzahl von Alternativen für den Straßennamen gewonnen. Vom Operator erfolgt eine Extraktionscodierung des zweiten Bestandteils der Adreßinformation, woraus ebenfalls eine Anzahl von Vorschlägen für den Straßennamen gewonnen wird. Bei der Konsistenzprüfung werden nur diejenigen Vorschläge nicht zurückgewiesen, bei denen die Ergebnisse dieser beiden Auswertungen untereinander kompatibel sind.

Im folgenden wird anhand von zwei Beispielen die obengenannte Ausführungsform genauer erläutert. Auf einer Sendung sei die Adresse

Bucklestr. 5

D-78457 Konstanz

statt der korrekten Adresse

Bücklestr. 5

D-78467 Konstanz

aufgetragen. In diesem Fall wird bei einer Extraktions-

codierung "78457 Buck 5" verwendet um aus einem Straßenverzeichnis die Eintragung "Buckley 5, Konstanz" zu erhalten. Dies ist eine korrekte Zuordnung der fehlerhaften Postleitzahl 78457. Bei der Konsistenzprüfung wird durch automatische Auswertung die Inkonsistenz von "Bucklestr." und Buckley" festgestellt und das betreffende Auswertungsergebnis zurückgewiesen. Wird andererseits die Adreßzeile mit dem ZIP-Code und der Ortsangabe mit einer hohen Fehlerrate OCR gelesen, während die Straßenangabe mit niedriger Fehlerrate gelesen wird, folgt vom Operator lediglich eine Eingabe des ZIP-Codes oder eine Anzahl von Buchstaben der Ortsangabe. Vorzugsweise wird in diesem Fall dem mit geringer Fehlerrate behafteten OCR-Ergebnis Vorrang gegenüber der Operatoreingabe gegeben.

Bei einer weiteren bevorzugten Ausführungsform der Erfindung erfolgt eine Auswertung der auf die Oberfläche von Sendungen aufgetragenen Angaben durch Videocodierung in den Fällen, in denen eine automatische Auswertung deshalb nicht erfolgreich war, weil Adreßangabe und Absenderangabe von der automatischen Auswertung nicht differenziert werden konnten. Dies tritt vor allem bei Post auf, bei der die Absenderangabe unmittelbar über oder unter der Adreßangabe aufgebracht ist, z. B. dänischer Post. Da beim Videocodieren eine Darstellung der Sendungsoberfläche für den Operator erfolgt, ist es üblicherweise leicht möglich, die betreffenden Informationen als Adreß- bzw. Absenderangabe zu identifizieren und gegebenenfalls eine Extraktionscodierung vorzunehmen.

Patentansprüche

1. Verfahren zur Bearbeitung von Sendungen in einem automatischen Adreßlesesystem, wobei ein Abbild der Adreßinformation aufweisenden Oberfläche jeder Sendung gewonnen und einer ersten automatischen Auswertung zugeführt wird, und nicht vollständig ausgewertete Adreßinformation einer ersten Video-Codierung zur Auswertung zugeführt wird, dadurch gekennzeichnet, daß die Adreßinformation derjenigen Abbilder, die bei der Video-Codierung nicht vollständig ausgewertet worden sind, unter Verwendung der Ergebnisse der Video-Codierung einer weiteren automatischen Auswertung zugeführt werden.
2. Verfahren zur Bearbeitung von Sendungen nach Anspruch 1, dadurch gekennzeichnet, daß die Adreßinformation diejenigen Abbilder, die bei der weiteren automatischen Auswertung nicht vollständig ausgewertet worden sind, einer weiteren Video-Codierung zur Auswertung unter Verwendung der Ergebnisse der weiteren automatischen Auswertung zugeführt werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß bei der ersten Video-Codierung eine Extraktionscodierung gemäß vorgegebener Extraktionsregeln erfolgt.
4. Verfahren nach Anspruch 2 oder 3, dadurch gekennzeichnet, daß bei der weiteren Video-Codierung eine Selektionscodierung derart erfolgt, daß aus einer Anzahl von alternativen Auswertungsergebnissen eine Selektion vorgenommen wird.
5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß die alternativen Auswertungsergebnisse aus den nicht eindeutigen Ergebnissen der Extraktionscodierungen gebildet werden.
6. Verfahren nach Anspruch 5, dadurch gekenn-

zeichnet, daß alternative Auswertungsergebnisse aus Adreßinformationen, die zusätzliche Sortierinformationen aufweisen, gebildet werden.

7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß eine Auswertung eines ersten Bestandteils der Adreßinformation erfolgt und daß eine Auswertung eines zweiten Bestandteils der Adreßinformation erfolgt und daß die Ergebnisse dieser Auswertungen in Hinblick auf gegenseitige Konsistenz geprüft werden.

8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß die Adreßinformation alle Abbilder die der ersten Video-Codierung zur Auswertung zugeführt worden sind, unter Verwendung der Ergebnisse der Video-Codierung der weiteren automatischen Auswertung zugeführt werden.

9. Verfahren nach Ansprüchen 1 bis 8, dadurch gekennzeichnet, daß die erste automatische Auswertung der Adreßinformation On-line oder Off-line erfolgt.

10. Verfahren nach Ansprüchen 1 bis 9, dadurch gekennzeichnet, daß Sendungen für die on-line keine vollständige weitere automatische Auswertung oder Auswertung durch weitere Video-Codierung der Adreßinformation erfolgt ist, mit einer Identifizierungsmarkierung (TID) für eine off-line vorzunehmende weitere automatische oder Video-Codierung versehen werden.

11. Verfahren nach Anspruch 1 bis 9, dadurch gekennzeichnet, daß zumindest bei einer der Video-Codierungen ein Vorschau-Codierungsverfahren verwendet wird.

12. Verfahren nach Anspruch 1 bis 11, dadurch gekennzeichnet, daß bei der Videocodierung eine Differenzierung zwischen Adreß- und Absenderinformation erfolgt.

13. Vorrichtung zur Durchführung des Verfahrens nach Ansprüchen 1 bis 12, mit einem automatisches Adreßlesesystem, das

- eine Vorrichtung zur Gewinnung von Abbildern von Sendungen,
- eine Vorrichtung zur automatischen Auswertung von Adreßinformation,
- eine Vorrichtung zur Video-Codierung von Adreßinformation,
- einen Imagecontroller zur Steuerung des Datenflusses zwischen der Vorrichtung zur automatischen Auswertung und der Vorrichtung zur Video-Codierung aufweist, wobei die Vorrichtung zur Video-Codierung eine Anzahl von Video-Codierplätzen aufweist,

dadurch gekennzeichnet, daß der Imagecontroller derart ausgebildet ist, daß von der Video-Codierung nicht vollständig ausgewertete Adreßinformation unter Verwendung der Ergebnisse der Video-Codierung der Vorrichtung zu einer weiteren automatischen Adreßauswertung zugeführt werden.

14. Vorrichtung nach Anspruch 13, dadurch gekennzeichnet, daß eine Vorrichtung vorgesehen ist, um Sendungen, deren Adreßinformationen nicht vollständig on-line ausgewertet worden ist, mit einer Identifizierungsinformation (TID) zu versehen.

Hierzu 3 Seite(n) Zeichnungen

- Leerseite -

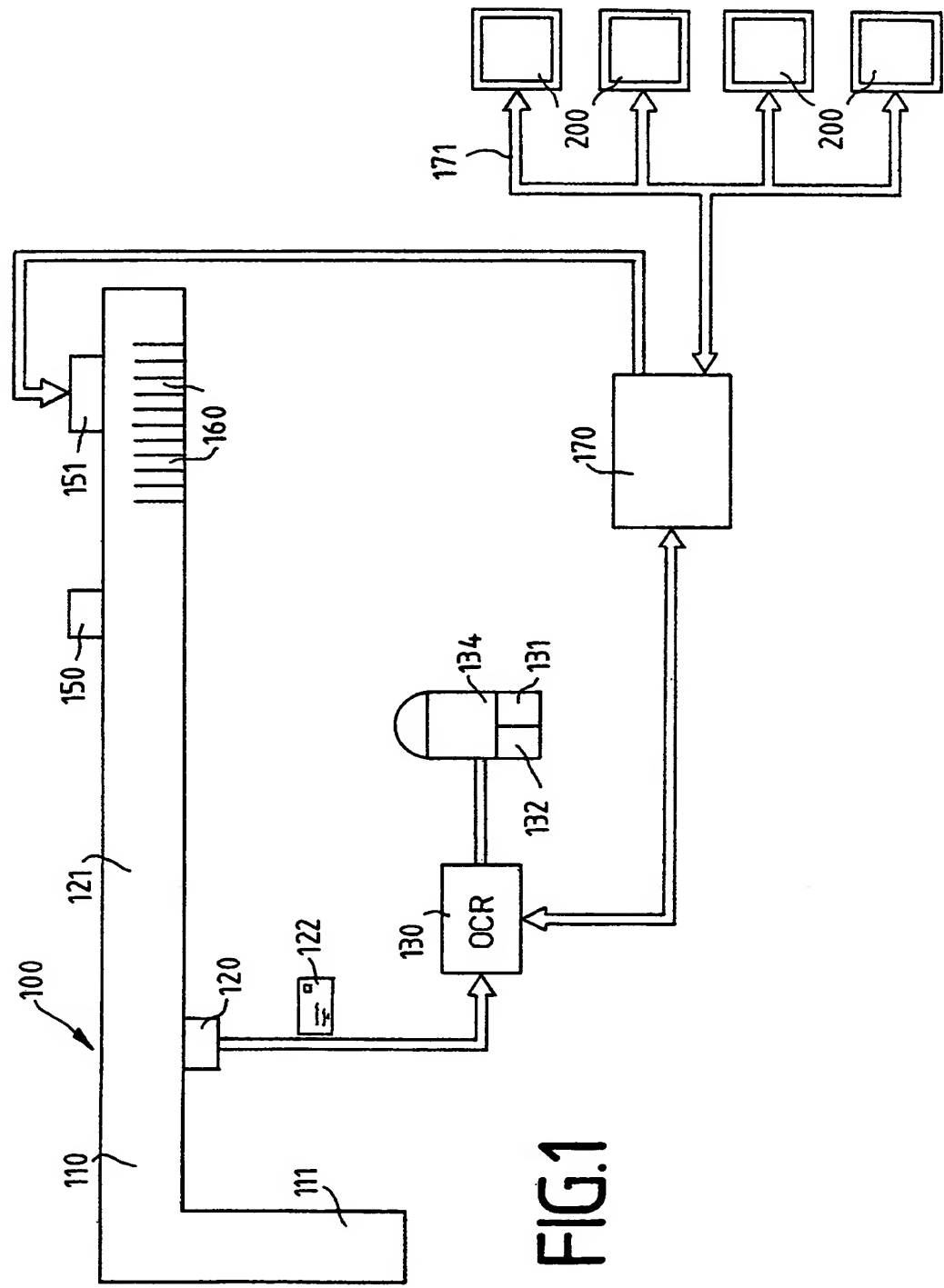


FIG. 1

FIG.2

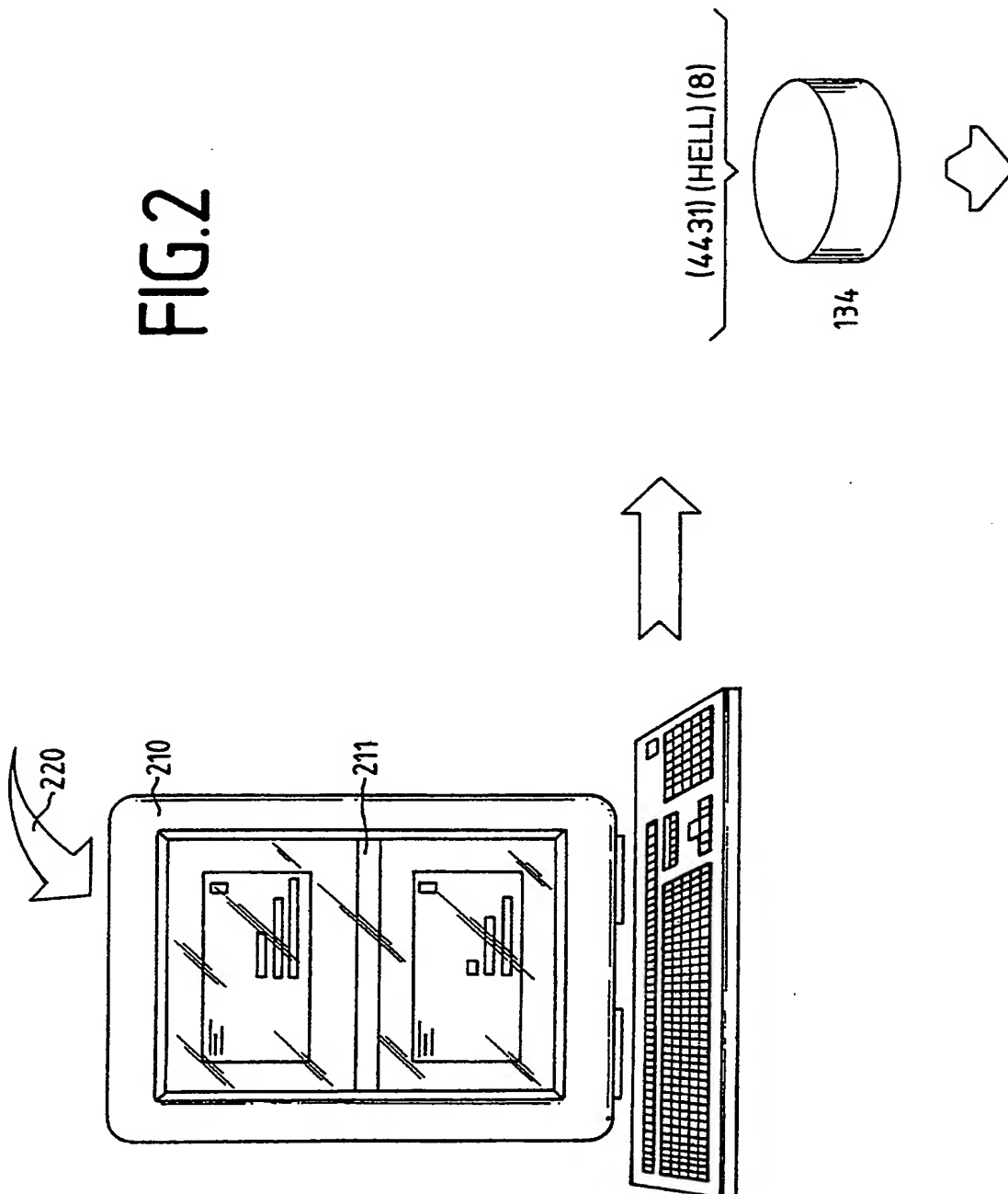


FIG. 3

